Multilingual phone recognition of spontaneous telephone

Corredor-Ardoy, C. Lamel, L. Adda-Decker, M. Gauvain, J.L.
Lab. d'Inf. pour la Mecanique et les Sci. de l'Ingenieur, CNRS, Orsay, France;

This paper appears in: Acoustics, Speech, and Signal Processing, 1998. ICASSP '98. 1998 IEEE International Conference on
Publication Date: 12-15 May 1998
Volume: 1
On page(s): 413 - 416 vol.1
Number of Pages: 6 vol. lxiii+3816
Meeting Date: 05/12/1998 - 05/15/1998
Location: Seattle, WA
ISSN: 1520-6149
ISBN: 0-7803-4428-6
INSPEC Accession Number:6033451
Digital Object Identifier: 10.1109/ICASSP.1998.674455
Posted online: 2002-08-06 21:40:28.0

Abstract
In this paper we report on experiments with phone recognition of spontaneous telephone s recognizers were trained and assessed on IDEAL, a multilingual corpus containing telepl French, British English, German and Castillan Spanish. We investigated the influence of th composition (size and linguistic content) on the recognition performance using context-ind hidden Markov models (HMMs) and phonotactic bigram models. We found that when tes speech data, using only spontaneous speech training data gave the highest phone accur languages, even though this data comprises only 14% of the available training data. The dependent (CD) HMMs reduced the phone error across the 4 languages, with the averag 51.9% from the 57.4% obtained with CI models. We suggest a straightforward way of dete phenomena. The basic idea is to remove sequences of consonants between two silence la recognized phone strings prior to scoring. This simple technique reduces the relative ave by 5.4%. The lowest phone error with CD models and filtering was obtained for Spanish (3 language average being 49.1%

Index Terms
Inspec
Controlled Indexing
hidden Markov models speech recognition telephony

Non-controlled Indexing
British English Castillan Spanish French German HMM IDEAL multilingual context-dependent hidden Markov models context-independent hidden Marko filtering linguistic content multilingual phone recognition non-speech phenom detection phonotactic bigram models relative average phone error rate sponta telephone speech training material composition
Author Keywords
Not Available

References

No references available on IEEE Xplore.

Citing Documents

No citing documents available on IEEE Xplore.

« View Search Results | Next Article »